

Pcache

Pcache is a small wrapper that can be used with any data move utility (dccp, lcg-cp, srmcp, ...).

It uses a limited amount of space on scratch device as a persistent local cache for Panda job input files.

This helps reduce load on storage systems caused by frequently-fetched files, such as DBRelease (but not just these).

It is particularly helpful on multi-core hosts.

Design goals

Works with any type of mass storage.

Requires no modifications to Panda or pilot.

Determines “hot” files automatically.

No extra administration – cache size is limited,
cleanup is automatic.

Files in cache can safely be deleted at any time.

Implementation

Syntax for copy command is assumed to be:

```
copy_cmd [-args] SRC DST
```

which is prepended with 'pcache.py':

```
pcache.py copy_cmd [-args] SRC DST
```

Pcache examines destination DST, if this is not a subdir of /scratch, command is executed “as-is”.

Implementation: caching

When the destination directory is under /scratch, (and pcache has not already cached the file), the file is copied into the pcache/CACHE directory instead.

A hard link is made from this file to the Panda job directory.

If file is already present in pcache, copying in is unnecessary: a hard link is made to the existing copy.

There is also handling for files 'in transit'.

Implementation: cleanup

At the end of each run, files are deleted from pcache directory (oldest first) until disk usage is below a specified limit.

Number of links is examined, and only files with nlink=1 are considered

Results

Running at UC since 9/2008

Very low # of errors due to pcache

Reduced rate of job failures due to input staging problems

~40% cache hit rate over long term

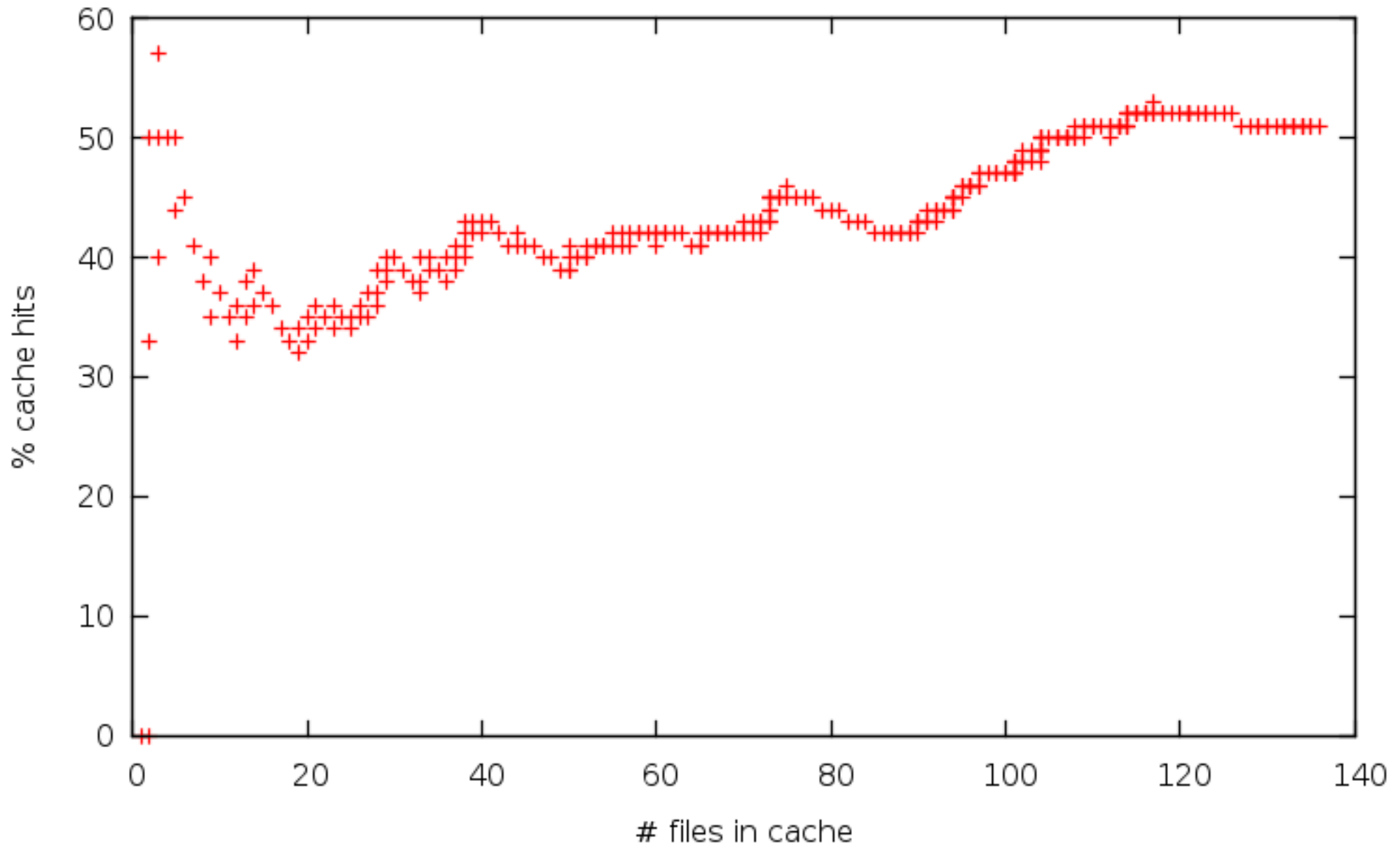
Statistics summary:

<http://www.mwt2.org/sys/pcache>

Most popular files

100743	DBRelease-5.5.1.tar.gz
56087	DBRelease-5.6.1.tar.gz
19256	DBRelease-6.3.1.1.tar.gz
17755	DBRelease-5.9.1.tar.gz
6249	EVNT.023234._00001.pool.root.1
3288	user08.AntonKapliy.uct3-edge5_20.lib._000025.lib.tgz
3061	DBRelease-6.1.1.tar.gz__DQ2-122590826
2864	EVNT.025872._00001.pool.root.1
2818	EVNT.023236._00001.pool.root
2444	EVNT.023237._00001.pool.root.1

Cache hit rate vs. cache size



Possible enhancements

Matching of Panda jobs to WN based on already-staged input files. Would require a way to query or catalog what is present in pcache (possibly via an extension of the LFC).

Links

Code and documentation are available at

<http://repo.mwt2.org/viewvc/pcache>

Please send comments or questions to:

cgw@hep.uchicago.edu